

Elicitation of Weibull priors

Nicolas Bousquet

EDF Research & Development

6 quai Watier 78401 Chatou - France

E-mail : nicolas.bousquet@edf.fr

Summary — Based on expert opinions, informative prior elicitation for the common Weibull lifetime distribution usually presents some difficulties since it requires to elicit a two-dimensional joint prior. We consider here a reliability framework where the available expert information states directly in terms of prior predictive values (lifetimes) and not parameter values, which are less intuitive. The novelty of our procedure is to weigh the expert information by the size m of a virtual sample yielding a similar information, the prior being seen as a reference posterior. Thus, the prior calibration by the Bayesian analyst, who has to moderate the subjective information with respect to the data information, is made simple. A main result is the full tractability of the prior under mild conditions, despite the conjugation issues encountered with the Weibull distribution. Besides, m is a practical focus point for discussion between analysts and experts, and a helpful parameter for leading sensitivity studies and reducing the potential imbalance in posterior selection between Bayesian Weibull models, which can be due to favoring arbitrarily a prior. The calibration of m is discussed and a real example is treated along the paper.

Key Words — subjective prior elicitation, Weibull distribution, expert opinion, virtual data, posterior prior.

1 Introduction

The versatile Weibull $\mathcal{W}(\eta, \beta)$ distribution, with density function

$$f_W(t|\eta, \beta) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \exp\left\{-\left(\frac{t}{\eta}\right)^\beta\right\} \mathbb{1}_{\{t \geq 0\}}$$

$(\eta, \beta) \in (0, \infty)^2$, is one of the most popular distributions in reliability and risk assessment (RRA) and many other fields, mainly chosen for modelling the lifetime T of an industrial system or component Σ [39]. In real-life studies, a Bayesian framework has often been highlighted when expert knowledge is available on Σ and observed lifetime data $\mathbf{t}_n = t_1, \dots, t_n$ are small-sized and possibly contain missing or censored values [2]. Such contexts are usually

encountered in industrial studies, especially when economical opportunities imply replacing a range of components at the same time, although they could have carried on running, and lead to “polluted” (e.g., censored) lifetime data. In those cases, all relevant sources of knowledge as expert opinion must be taken into account.

Therefore, numerous authors [34, 35, 6] have focused their work on the elicitation of a joint prior measure $\pi(\eta, \beta)$ that formalizes the expert knowledge, in order to integrate some decision-making function over the joint posterior distribution of these parameters, with density

$$\pi(\eta, \beta | \mathbf{t}_n) = \frac{\mathcal{L}(\mathbf{t}_n; \eta, \beta) \pi(\eta, \beta)}{\iint_{\mathbb{R}_+^2} \mathcal{L}(\mathbf{t}_n; \eta, \beta) \pi(\eta, \beta) d\eta d\beta},$$

where $\mathcal{L}(\mathbf{t}_n; \eta, \beta)$ denotes the data likelihood. There are two main difficulties using the Weibull distribution. First, its only conjugate prior distribution is continuous-discrete [36] and remains difficult to justify in real problems [20]. Second, the meanings of scale parameter η and shape parameter β greatly differ. Their values and correlation remain hard to assess by non-statistician experts, even though historical results [23] can be used to provide preferential values as a function of the behavior of the studied system. The methods proposed by the previous authors can suffer from this second defect and be applicable with difficulty. Therefore Kaminskiy & Krivtsov [20] recently provided a simple procedure to elicit a prior $\pi(\eta, \beta)$ using expert knowledge about the mean and standard deviation of the cumulative distribution function (cdf) F_W . They insisted on the fact that these values are easier to assess than parameter values.

However, especially in sensitive areas like nuclear safety [42], resorting to a Bayesian framework for reliability-based decision-helping implies defending the methodology of prior elicitation in front of control authorities, the traditional difficulty being the treatment of its subjective aspects [16]. According to most wishes expressed by decisionners in our practice, the elicitation of a “defendable” prior $\pi(\theta)$ (of any model parametrized by θ and not only Weibull) should respect the following items: (a) the quantity of subjective information can be directly compared, in terms of percentage, to the quantity of objective (data) information; and (b) the elicited prior must be unique.

One could add to this list other wishes as the practical handling of the prior (explicit features and easy sampling), which is of importance for sensitivity studies. Note that the first item requires to give a clear sense to the words “quantity of information”. Statistical definitions like inverse Fisher matrices or Shannon entropies are often too technical to be directly accessible to decisionners.

This article addresses those concerns. In the sequel, we consider an alternative elicitation of $\pi(\eta, \beta)$ defined as the reference posterior of virtual data of size m but calibrated from lifetime magnitudes directly given by experts, pursuing the worry of realism expressed by Kaminskiy & Krivtsov [20]. Since the virtual size corresponds to an intuitive measure of prior uncertainty, ratios of virtual and observed data sizes bring an understandable sense to the notion of “relative quantity of information”. It can appear simpler than standard deviations (or other typical statistical uncertainty measures) to discuss with non-statistician

experts and, especially, lead to more transparent choices to decisionners.

The structure of the paper is as follows. The full prior elicitation is detailed in (the largest) Section 2. This methodological section focuses on the calibration of hyperparameters, the aggregation of independent expert opinions and the equitability issues between Bayesian Weibull models. Posterior computation is considered in Section 3. A numerical application on a real case-study is treated along the paper to illustrate the methodology. A Discussion section ends the paper, presenting alternative results and some avenues for future research.

2 Prior elicitation

2.1 Principle

The central idea of prior elicitation comes from a simple vision of informative expert opinion already suggested by Lindley [26]. We suggest to consider that a perfect expert opinion should be, roughly speaking, similar to a real data survey, and provide an independent and identically distributed (i.i.d.) sample $\tilde{\mathbf{t}}_{\mathbf{m}} = (\tilde{t}_1, \dots, \tilde{t}_m)$ of lifetime data. Now, let π^J be a well-recognized formal representation of ignorance (namely, a noninformative prior), perceived as a reference benchmark measure on the parameter space. Assuming $\tilde{\mathbf{t}}_{\mathbf{m}}$ is known, the corresponding prior π should be the posterior distribution with density $\pi^J(\eta, \beta | \tilde{\mathbf{t}}_{\mathbf{m}})$.

Posterior priors. Priors built as virtual posteriors present some advantages in subjective Bayesian analysis. First, they are unique since only defined by π^J and the likelihood (often historically chosen in experiments). Second, the correlation between parameters is automatically assessed through the Bayes rule. Third, as said before, the ratio between the numbers of virtual and real data helps to yield an understandable answer to the (often unclear) question “what is the ratio between subjective and objective information” asked by cautious decision-makers. Finally, the aggregation of independent expert opinions is simply carried out through successive Bayes rules, the consensus virtual sample being an aggregation of all virtual data. This avoids choosing opinion pooling rules which can suffer from paradoxes [30].

Maybe the most famous of such elicited priors is Zellner’s g -prior [43]. Among others, Clarke [12], Neal [29], Kárny et al. [21], Lin et al. [25] and Morita et al. [28] examined various quantification of priors using virtual data. Kontkanen et al. [22] considered virtual data as practical tools for eliciting priors for Bayesian networks which may require automation in the treatment of their parameters.

However, because of various factors, especially subjective ones, the sample $\tilde{\mathbf{t}}_{\mathbf{m}}$ is not directly elicitable from an expert, and his or her information on lifetime T must be summarized through questioning processes [30]. This type of elicitation remains simple for distributions belonging to the natural exponential family, for which the resulting posterior priors are conjugate (cf. [32], § 5.3.3), because the virtual sample can be replaced by exhaustive statistics.

In the continuous Weibull case, unfortunately, the only exhaustive statistic is the full virtual likelihood. Therefore the questioning must be oriented such that it allows the calibration of nonexhaustive virtual statistics.

Expert questioning. In a concern of realism, following the ideas promoted by Kadane & Wolfson [19] and especially Percy [31] in the field of RRA, we consider that an expert is mainly capable of providing *observable* information on T , unconditionally to (η, β) . Indeed, the experts are usually not statisticians and should yield information independently from any parametrization choice (and even from any sampling model choice) made by a Bayesian analyst [21]. In other terms we assume that any realistic statistical summary of an expert opinion should be defined with respect to its associated prior predictive density, namely the prior density of plausible lifetime values

$$f_\pi(t) = \iint_{\mathbb{R}_+^2} f_W(t|\eta, \beta) \pi(\eta, \beta) d\eta d\beta. \quad (1)$$

The density form of de Finetti's representation theorem [14] is then invoked to ensure the unicity of priors elicited in this way, under mild conditions of exchangeability for sequences of values t_1, \dots, t_n, \dots . See Press [32], § 10.5, for more precisions. Linking typical magnitudes of the observable variable T with statistical specifications can be made through decision-theoretical arguments. In the following, we consider experts who can answer to a question similar to the following one:

Can you give estimates of relative costs (c_1, c_2) linked to two reliability-based decisions induced by the two mutually exclusive events $T \leq t_\alpha$ and $T > t_\alpha$?

provided t_α is given by the analyst (to improve the *phenomenon anchoring* of the expert and diminish subjective bias, cf. [41]) and denoting $\alpha = c_1/(c_1 + c_2) \in [0, 1]$. Doing so, following the standard criterion of decision theory, namely the *expected utility* [33], the analyst interprets t_α as the minimizer of the predictive Bayes risk

$$t_\alpha = \arg \min_{t_0 > 0} \int_0^\infty \Lambda(t_0, t|c_1, c_2) f_\pi(t) dt.$$

defined by some loss function $\Lambda(t_0, t|c_1, c_2)$ between the choice $T = t_0$ and the unknown truth $T = t$, inflicting c_1 to the event $t \leq t_0$ (underestimation) and c_2 to the contrary event $t > t_0$ (overestimation). The common choice

$$\Lambda(t_0, t|c_1, c_2) = |t - t_0| (c_1 \cdot 1_{\{t \leq t_0\}} + c_2 \cdot 1_{\{t > t_0\}}),$$

which underlies the analyst wants to penalize similarly small and large misestimations [33], leads to t_α taking the sense of the α -order prior predictive percentile

$$P_\pi(T < t_\alpha) = \int_0^{t_\alpha} f_\pi(t) dt = \alpha. \quad (2)$$

Therefore an alternative equivalent query is, perhaps simpler, *what is the risk α for Σ to break down before t_α ?*

In the following, we assume finally that for each available expert, a unique specified couple (t_α, α) among all elicitable can be considered as his or her *most trustworthy specification* (MTS) and must be exactly respected in the effective predictive prior modelling. Various reasons can be invoked for this. First, one cannot hope to elicit a prior $\pi(\eta, \beta)$ such that an arbitrary large number of specified couples (t_α, α) be exactly respected together, because of the limited flexibility of parametric distributions (Berger 1985 [4], chap. 3). Second, a MTS often appears as a reality since experts have usually more difficulty to speak in terms of extreme values rather than values close to the median behavior [30]. Typically, they can share a similar MTS while their extremes can differ (see Example 1). Other arguments can be related to decision-making: a cautious, *conservative* couple (t_α, α) can be favored by the analyst since the posterior analysis is focused on percentiles of higher, more critical orders.

EXAMPLE 1. *Table 1, already used in [8], summarizes two prior opinions about the lifetime T (in months) of a device belonging to the secondary water circuit of French nuclear plants. According to a large consensus in the RRA field, T is assumed to be well described by a Weibull distribution. Giving a normative sense to extreme events (90% credibility), these experts were not questioned at the same level of precision. \mathcal{E}_1 is a nuclear operator and spoke about a particular component, in terms of replacement costs. Conversely, \mathcal{E}_2 is a component vendor whose opinion took into account a variety of operating conditions. Costs invoked here were mainly related to mass production. Therefore the two experts can be considered independent. Hence the common median appears as a robust specification and is chosen as the MTS for both.*

	Credibility intervals (5%,95%)	Median value
expert \mathcal{E}_1	[200,300]	250
expert \mathcal{E}_2	[100,500]	250

Table 1: Expert opinions about the lifetime T of a nuclear device (in months).

2.2 A comfortable prior form

Let us choose π^J as the Jeffreys prior for Weibull. In more general Bayesian settings, Sun [37] proposed to favor the Berger-Bernardo reference prior [5] since it has slightly better properties of frequentist posterior coverage. But this prior requires at least $m \geq 2$ to get proper posteriors. This would be limiting in practice, when m is chosen small as it could be expected in cautious subjective assessments. Moreover, expert knowledge exerts here itself on T and not on any Weibull parametrization, therefore it seems relevant that a benchmark π^J be parametrization-invariant. See [12] for a straightforward defence of Jeffreys' prior in related problems, where a subjective posterior has to be compared in information-theoretic terms to an objective posterior. Thus we consider

$$\pi^J(\eta, \beta) \propto \eta^{-1} \mathbb{1}_{\{\eta \geq 0\}} \mathbb{1}_{\{\beta \geq \beta_0\}}$$

where $\beta_0 \geq 0$ is assumed to be fixed by objective reasons, like physical constraints. For instance, a reliability study focusing on industrial components submitted to aging leads to

choose $\beta_0 = 1$ as explained in Bacha (1998) [2], since involving a time-increasing failure rate. Without particular constraint, $\beta_0 = 0$. Scale invariance imposes no other lower bound than 0 for η .

Denote $\mathcal{GIG}(a, b, \gamma)$ the generalized inverse gamma distribution with density

$$f(x) = \frac{b^a \gamma}{\Gamma(a)} \frac{1}{x^{a\gamma+1}} \exp\left(-\frac{b}{x^\gamma}\right) \mathbb{1}_{\{x \geq 0\}}.$$

Reparametrizing x in $\mu = x^{-\gamma}$, $\mu \sim \mathcal{G}(a, b)$. Then our ideal prior is $\pi(\eta, \beta) = \pi(\eta|\beta)\pi(\beta)$, such that

$$\eta|\beta \sim \mathcal{GIG}(m, b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta), \beta), \quad (3)$$

$$\pi(\beta) \propto \frac{\beta^{m-1}}{b^m(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta)} \exp\left(m \frac{\beta}{b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta)}\right) \mathbb{1}_{\{\beta \geq \beta_0\}} \quad (4)$$

with $b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta) = \sum_{i=1}^m \tilde{t}_i^\beta$, and $\beta(\tilde{\mathbf{t}}_{\mathbf{m}}) = m(\sum_{i=1}^m \log \tilde{t}_i)^{-1}$. Both distributions are proper for all $m > 0$. The unknown virtual unsufficient statistics $b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta)$ and $\beta(\tilde{\mathbf{t}}_{\mathbf{m}})$ must be replaced in function of available expert information. The linkage between the prior form promoted in (3-4) and a MTS (t_α, α) elicitable for a given expert can be done as explained in the next proposition (proved in Appendix) and its corollary.

PROPOSITION 1. For $\alpha \in]0, 1[$ and $t_\alpha > 0$, define the function $b_\alpha : \mathcal{N}^* \times \mathcal{R}_+^*$ by

$$b_\alpha(m, \beta) = \left((1 - \alpha)^{-1/m} - 1\right)^{-1} t_\alpha^\beta. \quad (5)$$

Then $b_\alpha(m, \beta)$ is the only β -continuous function such that, being substituted to $b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta)$ in (3), Equation (2) is verified almost surely.

An immediate and pleasant consequence of replacing deterministic expression $b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta)$ by $b_\alpha(m, \beta)$ is that $\pi(\beta) \propto \beta^{m-1} \exp(-m\beta \log t_\alpha + m\beta/\beta(\tilde{\mathbf{t}}_{\mathbf{m}})) \mathbb{1}_{\{\eta \geq 0\}}$. We recognize here the general term of a gamma distribution truncated in β_0 . Finally the resulting prior is

$$\eta|\beta \sim \mathcal{GIG}(m, b_\alpha(m, \beta), \beta), \quad (6)$$

$$\beta \sim \mathcal{G}\left(m, \frac{m}{\tilde{\beta}(m)}\right) \mathbb{1}_{\{\beta \geq \beta_0\}} \quad (7)$$

where $\tilde{\beta}(m) = (\log t_\alpha - \beta^{-1}(\tilde{\mathbf{t}}_{\mathbf{m}}))^{-1}$. This result deserves some technical remarks.

- (i) The joint prior propriety imposes $\beta^{-1}(\tilde{\mathbf{t}}_{\mathbf{m}}) < \log t_\alpha$, namely $\prod_{i=1}^m \tilde{t}_i < t_\alpha^m$.
- (ii) The joint prior (6-7) can remain proper for all m extended on the half-line \mathcal{R}_+^* . Thus fuzzy or doubtful experts can be graded using $m \leq 1$. This might be valuable if a group of P experts is considered as yielding less information than P i.i.d. data, for instance because they are suspected of mutual influence.

- (iii) The $\mathcal{GIG}(a, b, \gamma)$ distribution was firstly used by Berger & Sun [6], b being assessed independently of β . However, this choice was made only because of the posterior conjugate properties conditionally to β (see § 3), and no meaning was given to the hyperparameters. Authors like Tsionas [39] adopted similar approaches.

2.3 Prior calibration

In addition to the MTS needed to define the prior form (6-7), supplementary prior information must be available for the calibration of $(m, \tilde{\beta}(m))$. In the two following paragraphs we consider some cases commonly encountered in RRA.

2.3.1 Calibrating $\tilde{\beta}(m)$

In RRA, it can occur that the analyst benefits from *qualitative information* on the nature of aging of Σ . For instance, assuming $\beta_0 = 0$, if the expert can answer the question *what is the probability $0 < \alpha_{\beta_e} < 1$ that Σ is submitted to aging?*, one would have a priori $P(\beta < \beta_e) = 1 - \alpha_{\beta_e}$ with $\beta_e = 1$ and consequently

$$\tilde{\beta}(m) = 2m\beta_e/\chi_{2m}^2(1 - \alpha_{\beta_e}) \quad (8)$$

where $\chi_m^2(q)$ is the q -order percentile of the χ_m^2 distribution. Other similar questions can be asked over accelerated aging ($\beta_e = 2$) and extreme cases ($\beta_e = 5$) reflecting inconceivable kinetics of aging in industrial applications [2, 23]. Otherwise, databases of typical β values (e.g., <http://www.barringer1.com/wdbase.htm>) can be used to quantify some alternative features of the gamma prior.

However, most frequently (as in Example 1), other *quantitative information* is available under the form of a single or several *credibility intervals*, one of whose bounds is the previously chosen MTS. We consider $p \geq 1$ supplementary (non-independent) specifications $\Omega_p = \{t_{\alpha_i}, \alpha_i\}_{i \in \{1, \dots, p\}}$, sorted by increasing order ($\alpha_i < \alpha_{i+1}$ and $t_{\alpha_i} < t_{\alpha_{i+1}}$). Given m , calibrating $\tilde{\beta}(m)$ under those predictive constraints can be done by minimizing a distance $\mathcal{D}_m(f^*, f_\pi)$ where f^* is a pdf of T respecting exactly the MTS and the specifications listed in Ω_p . To avoid dealing with the infinite number of possible f^* , we adopt the approach proposed by Cooke [13]: \mathcal{D} is chosen as the discrete Kullback-Leibler loss function between required and elicited marginal features

$$\begin{aligned} \mathcal{D}_m(f^*, f_\pi) &= \sum_{i=0}^p P_{f^*}(T \in [t_{\alpha_i}, t_{\alpha_{i+1}}]) \log \frac{P_{f^*}(T \in [t_{\alpha_i}, t_{\alpha_{i+1}}])}{P_{f_\pi}(T \in [t_{\alpha_i}, t_{\alpha_{i+1}}])} \\ &= \sum_{i=0}^p (\alpha_{i+1} - \alpha_i) \log \frac{(\alpha_{i+1} - \alpha_i)}{(\alpha_{i+1}^{(e)} - \alpha_i^{(e)})} \end{aligned} \quad (9)$$

where $t_{\alpha_0} = 0$, $t_{\alpha_{p+1}} = \infty$, $\alpha_0 = \alpha_0^{(e)} = 0$, $\alpha_{p+1} = \alpha_{p+1}^{(e)} = 1$, and for $i \in \{1, \dots, p\}$

$$\alpha_i^{(e)} = \iint F_W(t_{\alpha_i} | \eta, \beta) \pi(\eta, \beta) d\eta d\beta.$$

The convexity of this loss function in its argument π and, given m and β , the one-to-one continuous correspondence between $\pi(\beta|\tilde{\beta}(m))$ and $\tilde{\beta}(m)$ allows for a unique solution of the calibration problem

$$\tilde{\beta}^*(m) = \arg \min_{\pi(\cdot|\tilde{\beta}(m))} \mathcal{D}_m(f^*, f_\pi).$$

From (8), estimating $\tilde{\beta}^*(m)$ is similar to select $\alpha_{\beta_e} = 0.5$ and minimizing (9) in the prior median β_e^* . This provides a direct view of the underlying aging and numerical estimations were found slightly more robust than those of the prior mean, or those of the best order $\alpha_{\beta_e}^*$ if, conversely, β_e is fixed. Therefore we temporarily note $\tilde{\beta}(m) = \tilde{\beta}_m(\beta_e)$. For a given m , a combination of golden section search and successive parabolic interpolation [10] can achieve a robust optimization of β_e , provided the $\alpha_i^{(e)}$ are smoothly computed at each step of the algorithm. A smooth Monte Carlo estimation can be obtained using a unique importance sampling run $\beta_1, \dots, \beta_M \sim \mathcal{G}(m, m/\beta_{e,0})$, with large M , where $\beta_{e,0}$ is a chosen starting point:

$$\alpha_i^{(e)}(m, \beta_e) \simeq 1 - \left(\frac{\beta_{e,0}}{\tilde{\beta}_m(\beta_e)} \right)^m \frac{1}{M} \sum_{j=1}^M \left[\left(1 + \frac{t_{\alpha_i}^{\beta_j}}{b_\alpha(m, \beta_j)} \right) \exp \left(\beta_j \left[\frac{1}{\tilde{\beta}_m(\beta_e)} - \frac{1}{\beta_{e,0}} \right] \right) \right]^{-m}.$$

Note that

$$\text{Err}(m) = \mathcal{D}_m \left(f^*, f_{\pi(\cdot|\tilde{\beta}^*(m))} \right)$$

measures the expert incoherency with respect to the predictive Weibull distribution, given a virtual sample of size m in agreement with the expert opinion. If $\text{Err}(m)$ remains large for many m , the Weibull choice for the virtual data (and therefore for any real dataset, provided the expert is relevant for the problem) is at least debatable, not to say probably inappropriate.

2.3.2 Calibrating m

The calibration of m must be adapted to the experimental context. A decisionner can impose a given virtual size to improve the clarity of the posterior result. For instance, Marin and Robert [27] proposed to give to the virtual size parameter of Zellner's g -prior (on regressors of a gaussian linear regression problem) the value $m = 1$ by default. In a similar context, another possibility, proposed by Celeux et al. [11] and Liang et al. [24] among others, is to establish an upper hierarchical level in the Bayesian model by considering m as a random variable for which a weakly informative prior must be elicited. A last possibility is to use m as a discussion tool between the analyst and the expert, since the meaning of m is understandable outside the statistical field. Some heuristic methods in this sense are discussed in [8].

However our aim as a Bayesian analyst is mainly to measure the strenght of the expert opinion (assumed being correctly reflected through the prior modelling) through m . Besides, when the experts are no longer questionable and only a summary of their past opinions

remains available, it seems somewhat difficult to elicit a hyperprior on this parameter. Then we suggest that m should be integrated as the minimizer of the expert incoherency risk, namely

$$m^* = \arg \min_{m \geq 0} \text{Err}(m).$$

It is the analyst's decision to minimize this risk on \mathcal{I}^* (not to loose the virtual size meaning) or \mathcal{R}_+^* . In our experiments we chose \mathcal{R}_+^* to get the closest prior to the expert opinion. Obviously, one can avoid eliciting a too informative prior by limiting the minimization domain to $(0, n]$.

EXAMPLE 2. (PURSUING EXAMPLE 1). *For a continuum of values of m , we display on Figures 1 and 2 the optimized $\tilde{\beta}^*(m)$ and the corresponding risk $\text{Err}(m)$, respectively, for both experts. In both cases, the shape of $m \mapsto \text{Err}(m)$ allows for a unique solution m^* . We find $m^* = 3.36$ for expert \mathcal{E}_1 and $m^* = 2.50$ for expert \mathcal{E}_2 . This is logical since \mathcal{E}_1 is more informative than \mathcal{E}_2 . However, all values $\tilde{\beta}^*(m)$ for expert \mathcal{E}_1 appear unrealistic in an industrial physical context, testifying from exponential uncontrolled aging, assuming the Weibull model is correct. Especially, the calibrated $\tilde{\beta}^*(m^*) = 16.5$, which induces a peaked normal behavior of the prior predictive distribution (cf. [15]). On the contrary, the opinion of expert \mathcal{E}_2 remains physically plausible ($\tilde{\beta}^*(m^*) = 4.9$) although the underlying aging is still strong. The corresponding coverage matching error $|1 - (\alpha_{t_1}^{(e)} - \alpha_{t_2}^{(e)})/90\%|$, where $\alpha_{t_i}^{(e)}$ is the effective percentile order for $t_i \in \{100, 200\}$ or $t_i \in \{300, 500\}$, is plotted in Figure 3 and shows a good adequacy between the wanted and effective credible domains: less than 5% error in all cases, and less than 0.2% and 0.004% when choosing the calibrated m^* , for experts \mathcal{E}_1 and \mathcal{E}_2 respectively.*

2.4 Aggregation of independent expert opinions

In cases when the aggregation of $i = 1, \dots, p$ priors is chosen as a way to avoid interacting biases in a group of experts, it yields a similar information to that carried by a global virtual sample, which is the union of all experts' samples $\tilde{\mathbf{t}}_{\mathbf{m}_i}$. Because they are not explicitly known, one may use a concatenation of known samples $\tilde{\mathbf{s}}_{\mathbf{m}_1}, \dots, \tilde{\mathbf{s}}_{\mathbf{m}_p}$ from another model $\mathcal{M}(\eta, \beta)$ such that their parametric likelihood $(\eta, \beta) \mapsto \ell(\tilde{\mathbf{s}}_{\mathbf{m}_1}, \dots | \eta, \beta)$ leads to the same inference as the whole virtual sample. Indeed, we can show easily (cf. Lemma 1 in Appendix) that

$$\pi^J(\eta, \beta | \tilde{\mathbf{t}}_{\mathbf{m}_1}, \dots, \tilde{\mathbf{t}}_{\mathbf{m}_p}) = \pi_{\mathcal{M}}^J(\eta, \beta | \tilde{\mathbf{s}}_{\mathbf{m}_1}, \dots, \tilde{\mathbf{s}}_{\mathbf{m}_p}) \propto \pi^J(\eta, \beta) \prod_{i=1}^p \ell(\tilde{\mathbf{s}}_{\mathbf{m}_i} | \eta, \beta).$$

Next proposition, proved in Appendix, gives an example of such a likelihood (said *virtual likelihood*) for a single expert opinion.

PROPOSITION 2. *Consider $\tilde{\mathbf{s}}_{\mathbf{m}} = (k_{\alpha, m}, \beta_{t_{\alpha}, m})$, where $k_{\alpha, m} = ((1 - \alpha)^{-1/m} - 1)^{-1}$ and $\beta_{t_{\alpha}, m} = \tilde{\beta}(m)/(1 + \tilde{\beta}(m) \log t_{\alpha})$, as a sample whose components follow independently the $\mathcal{G}(m, m(t_{\alpha}/\eta)^{\beta})$ and $\mathcal{IG}(m, m\beta)$ distributions, respectively. Then it defines a virtual likelihood for an expert opinion summarized by $(t_{\alpha}, \alpha, \tilde{\beta}(m))$.*

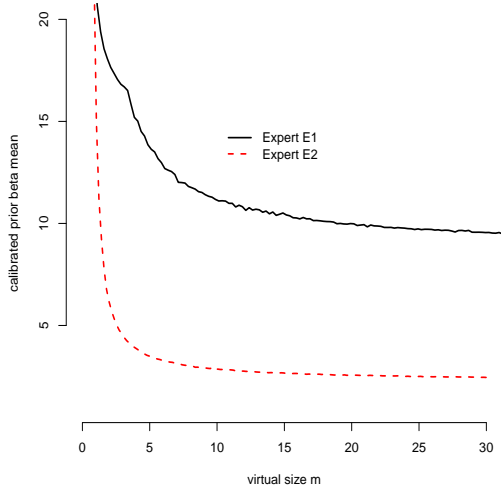


Figure 1: Calibrated prior mean $\tilde{\beta}^*(m)$ in function of m , for each expert opinion summarized in Table 1.

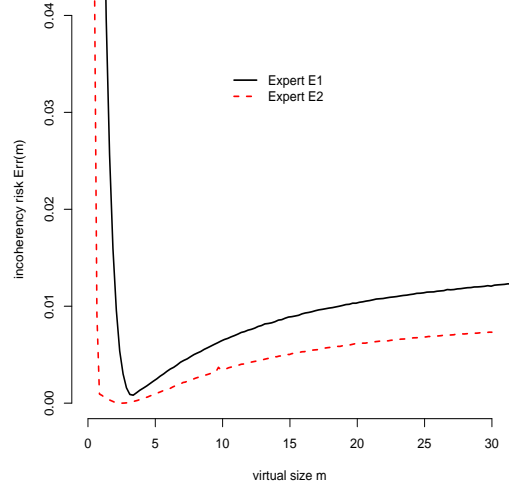


Figure 2: Expert incoherency risk $\text{Err}(m)$ in function of m , for each expert opinion summarized in Table 1. Minima of plots indicate calibrated values m^* .

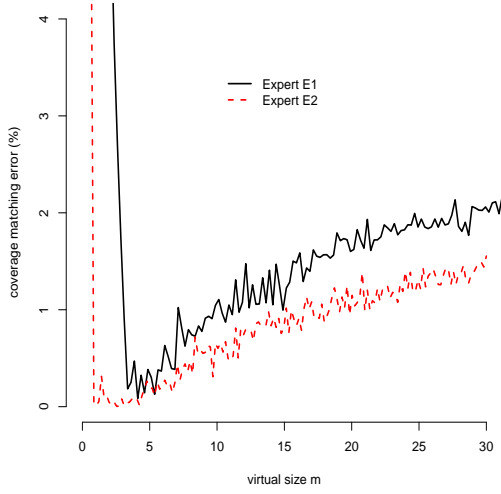


Figure 3: Prior predictive coverage matching error in percentage.

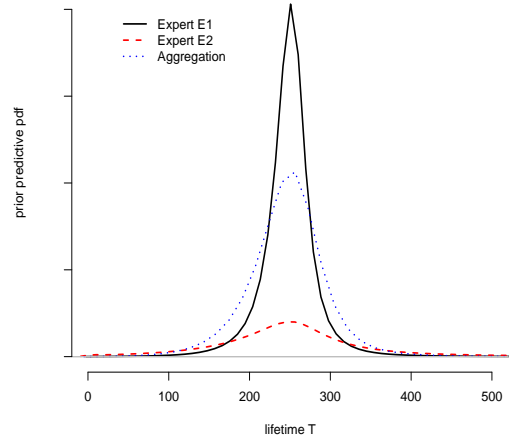


Figure 4: Prior predictive densities for both expert opinions and the aggregation of opinions.

After simple algebra, the resulting prior for all expert opinions is of the same form (6-7), for which (respecting intuition) $m = \sum_{i=1}^p m_i$, $b_\alpha(m, \beta) = \sum_{i=1}^p b_{\alpha_i}(m_i, \beta)$ and

$$\tilde{\beta}(m) = m \left(\sum_{i=1}^p \frac{m_i}{\tilde{\beta}_i(m_i)} \right)^{-1}.$$

EXAMPLE 3. (PURSUING EXAMPLE 2). Denote (π_1, π_2) the priors calibrated in Example 2 for each expert. Denote π_3 the aggregating prior. Although π_3 appears less relevant than π_2 with respect to a Weibull model in a RRA context, we have no supplementary information to weight the strenght of its corresponding virtual sample in π_3 . Then, by defect, π_3 is defined by

$$m = 5.86, \quad b_\alpha(\beta) = 7.33 \cdot (250)^\beta, \quad \tilde{\beta} = 8.30.$$

Corresponding prior predictive densities are plotted on Figure 4. As it could be expected, the aggregation prior realizes a trade-off between the two priors in the sense it favors the common median according to an intermediate peak, due to the addition of spread virtual data (expert \mathcal{E}_2) to concentrated virtual data (expert \mathcal{E}_1).

2.5 Prior equitability among Weibull models

Weibull models are often used as bricks for more general reliability models, like competing risk models [7] or mixtures [40]. Especially, a usual challenge in RRA to choose between exponential and Weibull models. Since the exponential is nested into the Weibull model ($\beta = 1$), a simple likelihood ratio test can be carried out in the frequentist framework. A Bayes factor is also easy to compute in our framework. Logically, both models share the same prior elicitation method, with the same MTS.

Then denote (m_E, m_W) the two corresponding virtual sizes, (π_E, π_W) the associated priors, and assume the more complex prior π_W has been calibrated. How should m_E be calibrated such that none of the prior Bayesian models is arbitrarily favored in absence of real data? The problem of defining such a *prior equitability* to reduce bias in posterior selection has been considered by many authors (see [11] for a review), who proposed several rules. Celeux et al. (2006) [11] gave decisive arguments to calibrate π_E such that it minimizes the Kullback-Leibler divergence between predictive distributions

$$m_E^* = \arg_{m_E} \min \text{KL}(f_{\pi_W}, f_{\pi_E}(\cdot|m_E))$$

where, after simple algebra,

$$f_{\pi_E}(t|m_E) = m_E \frac{b_\alpha^{m_E}(m_E, 1)}{(b_\alpha(m_E, 1) + t)^{m_E+1}}.$$

A unique Monte Carlo sampling of f_{π_W} can be used to get a smooth description of the KL divergence and its derivative in m , so that a coupled Newton-Raphson method can provide a good estimate of m_E^* . This strategy can be carried out on more complex Weibull models, sorting them through their decreasing order of degree of freedom. If all authors have emphasized the difficulty of this task when nested models are nonconjugate with multidimensional

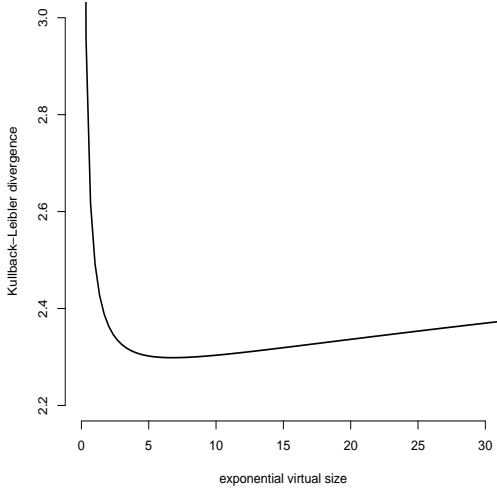


Figure 5: KL distance between prior predictive Weibull and exponential distributions dedicated to expert \mathcal{E}_2 's opinion, in function of the exponential virtual size m_E . The minimum is reached in m_E^* .

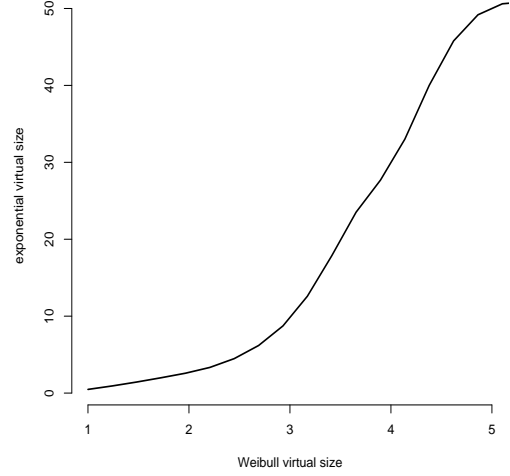


Figure 6: Correspondence between Weibull and exponential virtual sizes (m_W, m_E^*) for the expert \mathcal{E}_2 's opinion.

parameters, our framework leads to a rather simple optimization.

EXAMPLE 4. (PURSUING EXAMPLE 3). *For various values of m_E , the KL divergence is plotted for expert \mathcal{E}_2 on Figure 5. The KL convexity allows for a unique solution $m_E^* = 6.70$. When modifying m_W , the correspondence between m_W and m_E^* is plotted in Figure 6. A similar calculus for expert \mathcal{E}_1 leads to a very high value of m_E^* (upper than 200), because any exponential predictive distribution cannot approximate well a peaked normal distribution. An exponential assumption thus appears deeply irrelevant for this expert opinion.*

It could have been expected that $m_E^ \leq m_W$ since the more complex Weibull model should need more data than the exponential one to describe the same prior information. However, the model simplification reduces the global uncertainty in the effective prior predictive distribution. This has a direct impact on the virtual size m_E which varies inversely to uncertainty measures.*

3 Posterior inference

Thanks to the considerable development of numerical sampling methods, posterior computation is no longer burdensome in two-dimensional cases. Nonetheless, the conditional conjugation prior properties simplify the work of the Bayesian analyst. To be general in the RRA area and in relation with Example 5, we assume that observed data $\mathbf{t}_n = (t_1, \dots, t_n)$ contain r i.i.d

uncensored data $t_1^{(u)}, \dots, t_r^{(u)}$ and $n - r$ right-censored data. Denote $\delta_{\mathbf{t}_n}(\beta) = \sum_{i=1}^n t_i^\beta$ and $\beta_{\mathbf{t}_n} = r / \sum_{j=1}^r \log t_j^{(u)}$. Then the joint posterior distribution has density $\pi(\eta|\beta, \mathbf{t}_n)\pi(\beta|\mathbf{t}_n)$, such that

$$\begin{aligned}\eta|\beta, \mathbf{t}_n &\sim \mathcal{GIG}(m+r, b(m, \beta) + \delta_{\mathbf{t}_n}(\beta), \beta), \\ \pi(\beta|\mathbf{t}_n) &\propto \frac{b_\alpha^m(m, \beta)\beta^{m+r-1}}{(b_\alpha(m, \beta) + \delta_{\mathbf{t}_n}(\beta))^{m+r}} \exp\left\{-\beta\left(\frac{m}{\tilde{\beta}(m)} - \frac{r}{\beta_{\mathbf{t}_n}}\right)\right\} \mathbb{1}_{\{\beta \geq \beta_0\}}.\end{aligned}$$

It is enough to obtain approximate posterior sampling of β to get a complete joint sampling (using Gibbs sampling for η conditional to β). This can be made efficiently via the adaptive rejection sampling algorithm from Gilks and Wild [18]. In a noninformative context (i.e., when $m \xrightarrow{m>0} 0^+$) which can be easily adapted to a more general setting, Tsionas [39] proposed a gamma instrumental distribution $\rho(\beta)$ whose mean is calculated to optimize the acceptance rate.

A particular attention must be paid to the existence of posterior moments, especially the first one which defines the Mean Time To Failure (MTTF) in a RRA context. The conditional posterior mean of η is

$$\mathbb{E}[\eta|\beta, \mathbf{t}_n] = A(r, m, \beta) \left(m \cdot \eta_{e|\beta}^\beta + r \cdot \hat{\eta}_{|\beta}^\beta \right)^{1/\beta}$$

with $\eta_{e|\beta} = k_{\alpha, m}^{1/m} t_\alpha / m$, $\hat{\eta}_{|\beta} = r^{-1} \sum_{i=1}^n t_i^\beta$ the conditional MLE, and $A(r, m, \beta) = \Gamma(r+m-1/\beta)/\Gamma(r+m)$, so that the MTTF is not defined if $\beta_0 \leq 1/(m+r)$. More generally, using an important result of Sun and Speckman [38] (proof of Theorem 5), one can prove that the k^{th} moment of the posterior predictive density

$$\begin{aligned}\mathbb{E}[T^k|\mathbf{t}_n] &= \iint_{\mathbb{R}^2} \eta^k \Gamma(1+k/\beta) \pi(\eta, \beta|\mathbf{t}_n) d\eta d\beta, \\ &\propto \int_{\mathbb{R}_+} \Gamma(1+k/\beta) \Gamma(r+m-k/\beta) \pi(\beta|\mathbf{t}_n) d\beta,\end{aligned}$$

exists only if $\beta_0 > k/(r+m)$ for any $k > 0$. This result is especially useful in the sense it gives to the analyst a necessary requirement on the prior precision to justify the practical handling of the posterior predictive distribution through usual statistical summaries, in regards of the information available from really observed data. In mild conditions ($r \geq 5$ and $k \leq 2$), choosing a defect $\beta_0 = k/r$ appears as a practical calculus artifice, and the more justified choice $\beta_0 = 1$ in aging studies is sufficient to ensure in practice the existence of posterior predictive moments.

EXAMPLE 5. (PURSUING EXAMPLE 4). *We consider the right-censored lifetime data \mathbf{t}_n ($n = 18, r = 10$) from Table 2. They correspond to failure or stopping times collected on some similar devices \sum close to the one considered in Example 1. The maximum likelihood estimator (MLE) is $(\hat{\eta}_n, \hat{\beta}_n) = (140.8, 4.51)$ with estimated standard deviations $\hat{\sigma}_n = (7.3, 1.8)$. This strong aging is in agreement with the opinion of expert \mathcal{E}_2 .*

Choosing $\beta_0 = 0.1$ does not modify significantly the calibration of priors for both experts, and the resulting posterior distributions of the MTTF are plotted on Figure 7. The peak

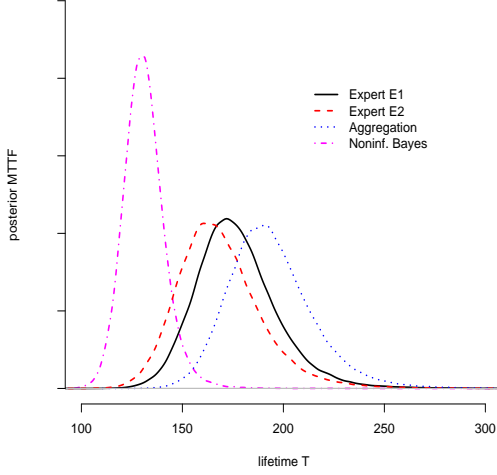


Figure 7: Posterior distributions of the MTTF for both experts and their aggregation, and in a noninformative framework.

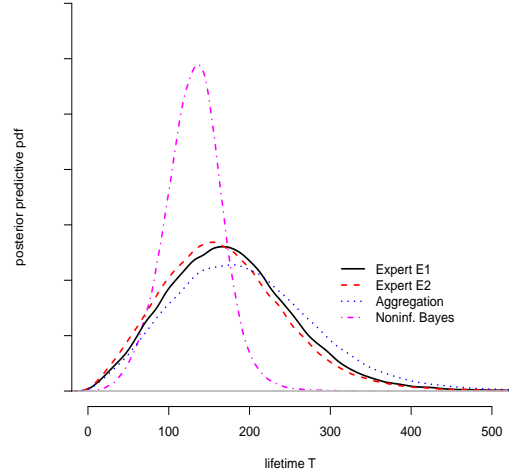


Figure 8: Posterior predictive densities for both experts and their aggregation, and in a noninformative framework.

observed when modelling a noninformative expert is due to the concentration of data far from time regions favored by the priors, and logically the aggregating prior, as the most informative, shifts the MTTF to the highest values.

The calibrated opinions of experts \mathcal{E}_1 and \mathcal{E}_2 have a relative weight of respectively 34% and 25% of the real data information transmitted to the posterior distribution, but, as it could be expected, their optimism in terms of lifetime has a strong influence on this important function of interest. Note however that due to variance increasing, the left tails of the posterior predictive pdf (Figure 8) are upper than the tails of a noninformative posterior predictive pdf. This means that, given a small t_γ (for instance a replacement time), the posterior estimation of $\gamma = P(T < t_\gamma)$ will be slightly overestimated - if we add the experts' opinions - with respect to this given by an only data-driven prediction. On the contrary, when t_γ is high, the noninformative posterior can appear somewhat too conservative.

real failure times	134.9, 152.1, 133.7, 114.8, 110.0, 129.0, 78.7, 72.8, 132.2, 91.8
right-censored times	70.0, 159.5, 98.5, 167.2, 66.8, 95.3, 80.9, 83.2

Table 2: Lifetimes (months) of nuclear components from secondary water circuits.

4 Discussion

The elicitation of a multidimensional prior, perceived as a reference posterior conditional to virtual data supposed to reflect a perfect expert opinion, is a practical way of assessing indirectly the correlations in the parameter space, coherently with the sampling model. Another important gain is the possibility of assessing the prior uncertainty in an understandable way by modulating the virtual size, for instance for sensitivity studies. This indicator of prior information might help to increase the trust of a decision-maker in the posterior beliefs and the acceptance of Bayesian assessments by control authorities. Unfortunately, such priors are often untractable since they require to assess nonexhaustive statistics of the virtual data. This is especially the case with the Weibull models.

In this article, however, we showed how this issue can be overcome, replacing those untractable statistics with functionals such that the resulting prior answers to the statistical specifications of the prior knowledge, under the form of percentiles. Note that other alternatives to percentiles could have been considered: for instance, following Percy [31], assume an expert can provide an estimate t_e of the marginal MTTF or the mode $\text{Md}[T]$, namely

$$\begin{aligned}\text{MTTF} &= \mathbb{E}[T] = \iint_{\mathbb{R}_+^2} \mathbb{E}[T|\eta, \beta] f_W(t|\eta, \beta) \pi(\eta, \beta) d\eta d\beta, \\ \text{Md}[T] &= \arg_{t>0} \max \iint_{\mathbb{R}_+^2} f_W(t|\eta, \beta) \pi(\eta, \beta) d\eta d\beta.\end{aligned}$$

The first estimate t_e is thus related to a quadratic loss function $\Lambda(t_0, t|c_1, c_2)$ which assumes the equality of costs $c_1 = c_2$ and a penalisation increasing with $|t_0 - t|$, while the second can be explained by the limit of a series of binary loss functions [33]. Following the principle described in Proposition 1, one should replace $b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta)$ in (3) by, respectively (see [8] for details),

$$\begin{aligned}b(m, \beta) &= \left(\frac{\Gamma(m)}{\Gamma(1 + 1/\beta)\Gamma(m - 1/\beta)} \right)^\beta t_e^\beta, \\ b(m, \beta) &= \left(\frac{m\beta + 1}{\beta - 1} \right) t_e^\beta.\end{aligned}$$

However, $\pi(\beta)$ is no longer (but remains close to) a gamma density, and furthermore these two specifications require conditions over m and the domain of variation of $\pi(\beta)$ to be usable. Indeed, one must guarantee $\beta_0 > 1$ to ensure $b(m, \beta)$ is well defined when t_e is specified as a unique mode. This is coherent with the Weibull features, since the Weibull distribution has a unique positive mode if and only if $\beta > 1$. As explained before, assuming aging is an equivalent prior constraint placed on the model.

Since assuming a prior predictive percentile can be provided by an oriented questioning, and fortunately leads to an explicit and versatile joint prior on Weibull parameters, we suggest Bayesian reliability analysts should favor, as much as possible, this kind of elicitation. This agrees with the vision historically promoted by Berger [4] (chap. 3) and Percy [31], who considered that quantile-based approaches pose among best elicitation methods, the estimation of probabilities of localization in given areas being simpler for experts than the

assessment of statistical moments.

Along the paper some remaining issues and limitations of the prior modelling have been evoked, which are now discussed as potential avenues for future researches. These researches, besides, will be dedicated to extend this methodology to other models which are often used in reliability studies, especially extreme value models whose links with Weibull distributions are well known.

It appears firstly that checking for the appropriateness of the Weibull model with respect to the virtual data is a crucial task, since it allows to reveal divergences between an expert opinion and the common sense of reliability practitioners when assuming the Weibull distribution for the lifetime of an industrial component submitted to aging. As the case for expert \mathcal{E}_1 illustrates, providing a small prior credibility interval can underlie unrealistic values for the Weibull parameters. This agrees with the well-known behavior of RRA experts of underestimating their self-uncertainty [23]. Therefore we suggest that spreading the prior credibility orders such that a qualitative requirement is reached (e.g., $P_\pi(\beta < 2) = 0.5$) can give a more reasonable summary of the real knowledge of Σ .

Another issue deals with the remaining uncertainty in expert opinion. In this article, we proposed a simple definition of the ratio of subjective and objective information based on virtual and real data sizes, understandable outside the community of statistics. It is clear however that information quantities updated through Bayesian inference are also strongly dependent on the possible *conflicting issues* between prior and real data, in the sense that both can favor regions of the sample and parameter spaces which are far from each other [17, 9]. Tools proposed in these two references should be carried out to check the internal coherency of the Bayesian model beforehand.

Besides, we did not consider here the remaining difficulties occurring when the expert can be suspected of bias, for instance because of *motivational reasons* [3] or dependency within a group. Many other tools, based on test experiments, have been proposed in the literature in order to quantify those biases (e.g. [34, 35, 23]). However, in absence of supplementary information, we preferred respecting the summarized expert opinions the best we could in our case-study.

Thus, the case of dependent experts has not been treated in this paper, since it remains controversial in Bayesian statistics [30] and probably deserves specific studies in the RRA area. Pursuing our view, two experts are dependent if they share virtual data from their past experience, or if a part of a virtual sample is produced dependently from the other sample. Dependency could then be introduced through a hierarchical mechanism of data production, in the spirit of the *supra-Bayesian* approaches promoted by Lindley [26]. This theme will also be considered in our next works.

In future studies, it will be necessary too to provide some calibration tools to take into account the expert uncertainty and bias. A first avenue can simply be to add a hierarchical level conditional to m . Indeed, we assumed here that the expert subjectivity mainly lies in the self-estimation of the costs associated with a reliability decision, and thus (using notations

from § 2.3) in the estimation of sorted orders $\alpha_1 < \dots < \alpha_p$. Let us denote $\tilde{\alpha}_1 < \dots < \tilde{\alpha}_p$ these prior estimates. If we pursue the virtual size idea, it appears logical to consider the α_i as correlated random variables such that, a priori,

$$\alpha_1, \alpha_2 - \alpha_1, \dots, \alpha_p - \alpha_{p-1}, 1 - \alpha_p \sim \mathcal{D}_{ir}(\nu_1, \dots, \nu_{p+1})$$

with $\nu_i - 1$ being the number of virtual “past” observations of the event $t_{\alpha_{i-1}} \leq T \leq t_{\alpha_i}$, which induces $\sum_{i=1}^{p+1} \nu_i = m + p + 1$, the Dirichlet distribution appearing naturally from well-known conjugation properties. Imposing $E[\alpha_i - \alpha_{i-1}] = \tilde{\alpha}_i - \tilde{\alpha}_{i-1}$ leads to elicit

$$\nu_i = (m + p + 1)(\tilde{\alpha}_i - \tilde{\alpha}_{i-1}).$$

The obvious correlation between prior estimates $\tilde{\alpha}_i$ threatens to underestimate the prior uncertainty of the α_i , so that it appears more appropriate to help the expert providing *conditional* probabilities by answering to the following question: *what is the risk $\alpha_i - \alpha_{i-1}$ for Σ to break down before t_{α_i} knowing Σ still runs after $t_{\alpha_{i-1}}$?* Thus, the randomization of values $(\alpha_1, \dots, \alpha_p)$ imposes indirectly a prior distribution $\pi(\tilde{\beta}^*(m))$ rather than a single value, and consequently, m should now be calibrated as the minimizer of the *expected* expert incoherency risk

$$m^* = \arg \min_{m \geq 0} E_{\pi(\tilde{\beta}^*(m))} \left[\mathcal{D}_m \left(f^*, f_{\pi(\cdot|\tilde{\beta}^*(m))} \right) \right].$$

But this calibration remains in facts difficult to carry out, since the untractability of $\pi(\tilde{\beta}^*(m))$ imposes a double Monte Carlo approximation coupled to an optimization strategy. Our next research will focus on simplifying this computational work.

Finally, we remind to the reader that other elicitation approaches of percentile orders are possible, mainly based on the establishment of expert preferences on a series of bettings such that α is not directly estimated but is progressively bounded [30]. Such methods are robust with respect to perturbations of the often criticized expected utility criterion, since they can lead to results that are independent of the behavior of the expert face to his or her self-perception of the risk [1]. It thus should be worthy to adapt our modelling and the calibration aspects to this type of elicitation, which could lead to more cautious, credible statistical features of the prior modelling.

Acknowledgements

The author thanks Gilles Celeux (INRIA), Eric Parent and Merlin Keller (ENGREF) for numerous enriching discussions, advices and references. He thanks Francois Billy, Emmanuel Remy, Alberto Pasanisi (EDF R&D) too for fruitful discussions about the specific issues raised by the industrial context of the study.

References

- [1] Abdellaoui, M. (2000). Parameter-free elicitation of utilities and probability weighting functions, *Management Science*, 46: 1497-1512.

- [2] Bacha, M., Celeux, G., Idée, E., Lannoy, A. and Vasseur, D. (1998). *Estimation de modèles de durées de vie fortement censurées*, Eyrolles (in French).
- [3] Benson, P.G., and Nichols, M.L. (1982). An investigation of motivational bias in subjective predictive probability distribution. *Decision Sci.*, **13**: 10-59.
- [4] Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.
- [5] Berger, J.O and Bernardo, J.M. (1992). On the development of reference priors (with discussion). In: J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, Eds., *Bayesian Statistics 4*, Oxford University Press: 35-60.
- [6] Berger, J.O. and Sun, D. (1993). Bayesian analysis for the Poly-Weibull Distribution, *J. Amer. Statist. Assoc.*, 88: 1412-1418.
- [7] Bertholon, H., Bousquet, N., Celeux, G. (2006). An alternative competing risk model to the Weibull distribution in lifetime data analysis, *Lifetime Data Analysis*, 12: 481-504.
- [8] Bousquet, N. (2006). A Bayesian analysis of industrial lifetime data with Weibull distributions, Research Report RR-6025, INRIA.
- [9] Bousquet, N. (2006). Diagnostics of prior-data agreement in applied Bayesian analysis. *J. Appl. Statist.*, 35: 1011-1029.
- [10] Brent, R. (1973). *Algorithms for Minimization without Derivatives*. Englewood Cliffs N.J.: Prentice-Hall.
- [11] Celeux, G., Marin, J.M., Robert, C.P. (2006). Sélection bayésienne de variables en régression linéaire. *Journal de la Société Française de Statistique*, 147: 59-79 (in French).
- [12] Clarke, B.S. (1996). Implications of reference priors for prior information and for sample size, *J. Amer. Statist. Assoc.*, 91: 173-184.
- [13] Cooke, R.M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. New York: Oxford University Press.
- [14] De Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré* (in French), 7: 1-68.
- [15] Dodson, B. (2006). *The Weibull Analysis Handbook (2nd ed.)*. ASQ Quality Press, p. 7.
- [16] Efron, B. (1986). Why isn't everyone a Bayesian? (with Discussion), *The American Statistician*, 40: 1-11.
- [17] Evans, M., & Moshonov, H. (2006). Checking for prior-data conflict. *Bayesian analysis*, 1: 893-914.
- [18] Gilks, W.R., Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics*, 41: 337-348.
- [19] Kadane, J.B., Wolfson, J.A. (1998). Experiences in elicitation, *The Statistician*, 47: 3-19.
- [20] Kaminskiy, M.P., and Krivtsov, V.V. (2005). A Simple Procedure for Bayesian Estimation of the Weibull Distribution, *IEEE Trans. Reliability*, 54: 612-616.

- [21] Kárný, M., Nedoma, P., Khailova, N., Pavelková, L. (2003). Prior information in structure estimation. *IEEE Proceedings in Control Theory and Applications*, 150: 643-653.
- [22] Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., Grünwald, P. (1998). Bayesian and information-theoretic priors for Bayesian networks parameters. *Lecture notes in computer science*. In the proceedings of the ECML-98 European conference on machine learning, Chemnitz, Germany, 1398: 89-94.
- [23] Lannoy, A. and Procaccia, H. (2001). *L'utilisation du jugement d'expert en sûreté de fonctionnement*, Tec & Doc (in French).
- [24] Liang, F., Paulo, R., Molina, G., Clyde, M., Berger, J. (2008). Mixtures of g-priors for Bayesian variable selection. *J. American Statist. Assoc.*, 103: 410-423.
- [25] Lin, X., Pittman, J. and Clarke, B. (2007). Information Conversion, Effective Samples, and Parameter Size. *IEEE Trans. Info. Theory*, 53: 4438-4456.
- [26] Lindley, D.V. (1983). Reconciliation of probability distributions. *Operations Research*, 31: 866-880.
- [27] Marin, J.M., Robert, C.P. (2010). Les bases de la statistique bayésienne. Rapport des Universités Montpellier II & Dauphine - CREST (in French).
- [28] Morita, S., Thall, P.F. and Mueller, P. (2007). Determining the effective sample size of a parametric prior. *UT MD Anderson Cancer Center Department of Biostatistics*, Working Paper Series. Working Paper 36.
- [29] Neal, R.M. (2001). Transferring prior information between models using imaginary data. Technical Report 0108, Dept. Statistics, Univ. Toronto.
- [30] O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., Rakow, T. (2006), *Uncertain Judgements: Eliciting Expert Probabilities*. John Wiley and Sons, Chichester.
- [31] Percy, D.F. (2003). Subjective Reliability Analysis Using Predictive Elicitation. In: *Mathematical and Statistical Methods in Reliability*, B.H. Lindqvist & K.A. Doksum (eds). *Quality, Reliability & Engineering Statistics 7*, World Scientific Publishing Co.: Singapore, pp. 57-72.
- [32] Press, S.J. (2003). *Subjective and Objective Bayesian Statistics* (second edition), New York: Wiley.
- [33] Robert, C.P. (2001). *The Bayesian Choice. A Decision-Theoretic Motivation* (second edition), Springer.
- [34] Singpurwalla, N.D. and Song, M.S. (1986). An analysis of Weibull lifetime data incorporating expert opinion, in *Probability and Bayesian Statistics* (R.Viertl ed.), Plenum Pub.Corp.: 431-442.
- [35] Singpurwalla, N.D. (1988). An interactive PC-Based procedure for reliability assessment incorporating expert opinion and survival data, *J. Amer. Statis. Assoc.*, 83: 43-51.
- [36] Soland, R. (1969). Bayesian analysis of the Weibull process with unknown scale and shape parameters, *IEEE Transactions on Reliability*, 18: 181-184.
- [37] Sun, D. (1997). A note on noninformative priors for Weibull distributions, *J. Statist. Planning and Inference*, 61: 319-338.

- [38] Sun, D., Speckman, P.L. (2005). A note on the nonexistence of posterior moments, *Can. J. Statist.*, 33: 591-601.
- [39] Tsionas, E.G. (2000). Posterior analysis, prediction and reliability in three-parameter Weibull distributions, *Commun. Statist. - Theory Meth.*, 29(7): 1435-1449.
- [40] Tsionas, E.G. (2002). Bayesian analysis of finite mixtures of distributions, *Commun. Statist. - Theory Meth.*, 31(1): 37-48.
- [41] Tversky, A., Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases, *Science*, 185: 1124-1131.
- [42] Unwin, S.D., Cazzoli, E.G., Davis, R.E., Khatib-Rahbar, M., Lee, M., Nourbakhsh, H., Park, C.K., Schmidt, E. (1989). An information-theoretic basis for uncertainty analysis: application to the QUASAR severe accident study, *Reliability Engineering & System Safety*, 26: 143-162.
- [43] Zellner, A. (1986). On assessing Prior Distributions and Bayesian Regression analysis with g-prior distribution regression using Bayesian variable selection, In *Bayesian inference and decision techniques : Essays in Honor of Bruno De Finetti*: 233-243, North-Holland, Elsevier.

Appendix

Proof of Proposition 1. Denote $f_m(\beta) = \frac{\beta^{m-1}}{(b_\alpha(m, \beta))^m} \exp\left(m \frac{\beta}{\beta(\tilde{\mathbf{t}}_{\mathbf{m}})}\right)$ and $g_m(\beta) = \left(1 + \frac{t_\alpha^\beta}{b_\alpha(m, \beta)}\right)^{-m}$. Then, renaming $b(\tilde{\mathbf{t}}_{\mathbf{m}}, \beta)$ in $b_\alpha(m, \beta)$ in (3-4),

$$\begin{aligned} P_\pi(T \leq t_\alpha) &= \iint_{\mathbb{R}_+ \times \mathbb{R}_+} F_W(t_\alpha | \eta, \beta) \pi(\eta | \beta) \pi(\beta) d\eta d\beta, \\ &= 1 - \Delta_m \int_{\mathbb{R}_+} f_m(\beta) g_m(\beta) d\beta \end{aligned}$$

where $\Delta_m^{-1} = \int_{\mathbb{R}_+} f_m(\beta) d\beta$ which must be necessarily finite to get a proper $\pi(\beta)$. Assuming $P_\pi(T \leq t_\alpha) = \alpha$ leads to $\int_{\mathbb{R}_+} f_m(\beta) h_m(\beta) d\beta = 0$ where $h_m(\beta) = g_m(\beta) - (1 - \alpha)$ which is continuous in $\beta \geq 0$ if $\beta \mapsto b_\alpha(m, \beta)$ is assumed continuous for any $m > 0$. Since $\beta \mapsto f_m(\beta) > 0$ except possibly on a finite number of points in \mathbb{R}_+ , $h_m(\beta) = 0$ almost everywhere in \mathbb{R}_+ . Expression (5) follows immediately.

LEMMA 1. Denote \mathcal{M}_1 and \mathcal{M}_2 two sampling models with same parameter θ . Denote $\tilde{\mathbf{t}}_{\mathbf{m}_1}$ a non-observed \mathcal{M}_1 -sample with likelihood $\ell_{\mathcal{M}(1)}$. Let $\tilde{\mathbf{s}}_{\mathbf{m}_1} = g(\tilde{\mathbf{t}}_{\mathbf{m}_1})$ be an observed \mathcal{M}_2 -sample with likelihood $\ell_{\mathcal{M}(2)}$, such that $\theta \mapsto \ell_{\mathcal{M}(2)}(\theta) \propto \ell_{\mathcal{M}_i^{(1)}}(\theta)$. Denote π^J a prior measure on θ and $\pi_{\mathcal{M}(j)}^J$ its posterior knowing likelihood $\ell_{\mathcal{M}(j)}$. Then for p various samples $\tilde{\mathbf{t}}_{\mathbf{m}_1}, \dots, \tilde{\mathbf{t}}_{\mathbf{m}_p}$, one has

$$\pi_{\mathcal{M}(1)}^J(\theta | \tilde{\mathbf{t}}_{\mathbf{m}_1}, \dots, \tilde{\mathbf{t}}_{\mathbf{m}_p}) = \pi_{\mathcal{M}(2)}^J(\theta | \tilde{\mathbf{s}}_{\mathbf{m}_1}, \dots, \tilde{\mathbf{s}}_{\mathbf{m}_p})$$

Proof. The proof is straightforward, and we can consider only two non-observed samples $\tilde{\mathbf{t}}_{\mathbf{m}_1}$ and $\tilde{\mathbf{t}}_{\mathbf{m}_2}$ (corresponding possibly to two independent expert opinions). Then

$$\begin{aligned}
\pi_{\mathcal{M}^{(1)}}^J(\theta|\tilde{\mathbf{t}}_{\mathbf{m}_1}, \tilde{\mathbf{t}}_{\mathbf{m}_2}) &\propto \ell_{\mathcal{M}^{(1)}}(\tilde{\mathbf{t}}_{\mathbf{m}_1}|\theta) \pi_{\mathcal{M}^{(1)}}^J(\theta|\tilde{\mathbf{t}}_{\mathbf{m}_2}), \\
&\propto \ell_{\mathcal{M}^{(2)}}(\tilde{\mathbf{s}}_{\mathbf{m}_1}|\theta) \pi^J(\theta) \ell_{\mathcal{M}^{(1)}}(\tilde{\mathbf{t}}_{\mathbf{m}_2}|\theta), \\
&\propto \ell_{\mathcal{M}^{(2)}}(\tilde{\mathbf{s}}_{\mathbf{m}_1}|\theta) \ell_{\mathcal{M}^{(2)}}(\tilde{\mathbf{s}}_{\mathbf{m}_2}|\theta) \pi^J(\theta), \\
&\propto \ell_{\mathcal{M}^{(2)}}(\tilde{\mathbf{s}}_{\mathbf{m}_1}|\theta) \pi_{\mathcal{M}^{(2)}}^J(\theta|\tilde{\mathbf{s}}_{\mathbf{m}_2}), \\
&\propto \pi_{\mathcal{M}^{(2)}}^J(\theta|\tilde{\mathbf{s}}_{\mathbf{m}_1}, \tilde{\mathbf{s}}_{\mathbf{m}_2})
\end{aligned}$$

and the equality follows by immediate normalization.

Proof of Proposition 2. The Weibull likelihood of the m virtual data summarized by $(t_\alpha, \alpha, \tilde{\beta}(m))$ is proportional to $\ell(\eta, \beta) = \pi(\eta, \beta)/\pi^J(\eta, \beta)$ where $\pi(\eta, \beta)$ is defined by (6-7). Modify the parametrization using $\mu = \eta^{-\beta}$. Thus

$$\begin{aligned}
\ell(\mu, \beta) &\propto \mu^m \exp(-\mu b_\alpha(m, \beta)) \beta^m \exp\left(-m \frac{\beta}{\tilde{\beta}(m)}\right), \\
&\propto \left[\left\{\mu(t_\alpha)^\beta\right\}^m \exp\left\{-\mu k_{\alpha,m}(t_\alpha)^\beta\right\}\right] \left[\beta^m \exp\left\{-m \frac{\beta}{\tilde{\beta}(m)}\right\} \exp\{-m\beta \log t_\alpha\}\right], \\
&\propto \ell_1(k_{\alpha,m}|\mu, \beta) \ell_2(\beta_{t_\alpha,m}|\beta)
\end{aligned}$$

where

$$\ell_1(k_{\alpha,m}|\mu, \beta) \propto \frac{\left\{\mu(t_\alpha)^\beta\right\}^m}{\Gamma(m)} k_{\alpha,m}^{m-1} \exp\left\{-k_{\alpha,m}(t_\alpha)^\beta\right\}$$

which is the likelihood arising from considering $k_{\alpha,m} \sim \mathcal{G}(m, m\mu(t_\alpha)^\beta)$, and

$$\ell_2(\beta_{t_\alpha,m}|\beta) \propto \frac{(m\beta)^m}{\Gamma(m)} \beta_{t_\alpha,m}^{m-1} \exp\left\{-m\beta \frac{1 + \tilde{\beta}(m) \log t_\alpha}{\tilde{\beta}(m)}\right\}$$

which is the likelihood arising from considering $\beta_{t_\alpha,\delta} \sim \mathcal{IG}(m, m\beta)$.